An Examination of the Statistical Discrepancy and Private Investment Expenditure

Christopher Bajada

ISSN: 1036-7373
An Examination of the Statistical Discrepancy and Private Investment Expenditure

Christopher Bajada

School of Finance and Economics
University of Technology, Sydney
Broadway, NSW 2007

Abstract

The statistical discrepancy is often used to gauge the reliability of national accounts data. Particularly since the mid-1980's the statistical discrepancy in Australia has grown significantly in size and variance. In this paper we demonstrate that the overwhelming contribution to the size of the statistical discrepancy is mismeasurement of private investment expenditure. We demonstrate that this mismeasurement not only adds to the volatility of investment but may have a significant impact on the volatility of the business cycle in general.

Keywords: statistical discrepancy, national accounts, investment, business cycles.

JEL Classification: E32, C82

*I would like to thank Ross Milbourne and Graham Voss for their comments and suggestions. Naturally, any errors and omissions are my own.
I Introduction

In the Australian National Accounts (ANA), the Australian Bureau of Statistics (ABS) estimates the size of economic activity in Australia by calculating Gross Domestic Product (GDP). There are three alternative measures by which the ABS calculates GDP:\(^1\): (1) the expenditure approach; (2) the income approach; and (3) the production approach. In principle these three methods should yield the same results but in practice they do not. The ABS statistician is required to introduce a "Statistical Discrepancy" item in the ANA in order to reconcile the income side with the expenditure side.

The statistical discrepancy in the national accounts has in recent years increased significantly both in mean and variance. Since 1970 the discrepancy has averaged 2% of GDP and in more than 36% of quarters the growth in the statistical discrepancy was greater than or equal to the growth in GDP. Unlike most of the OECD countries, Australia's statistical discrepancy is quite large and it has been so particularly since the mid-1970's. In recent years, particularly since the mid-1980's, the statistical discrepancy has been predominantly positive and growing. Evidently the implication of this are that one or more of the major expenditure components in the ANA, such as investment and consumption, are under-reported. Clearly there is an issue regarding accuracy and reliability.

The question of how accurate and reliable are the national accounts is important for many reasons. It is of considerable policy interest to have accurately measured economic data because these are intended to provide not only a comprehensive and systematic summary of economic activity, but also a resource from which to gauge economic policies. Secondly, the existence of a non-negligible and volatile statistical discrepancy has implications not only for investigating economic theories but

---

\(^1\) Recently the ABS implemented the System of National Accounts (SNA93) into the ANA. While significantly contributing to an improvement in the measurement of national output, the changes have only had a small impact on the movement of GDP [ABS (1998)]. The three alternative measures of calculating GDP are no longer explicitly published, replaced now by a single measure. To ensure that the components do balance, the statistical discrepancy is now allocated to each of the components based on information from input-output tables. Although the statistical discrepancy is no longer explicitly reported, an estimate of its size is nevertheless possible to construct.
implications also for the business cycle in general. Consequently it is important to know something about the statistical properties of the discrepancy.

Weale (1992) proposed a maximum likelihood procedure to identify whether income or expenditure measures of GDP contribute most to the size of the statistical discrepancy. The aim of this paper is twofold. First, to generalise the procedure in order to determine which component(s) of the national accounts have contributed most to the statistical discrepancy. Second, to investigate the implication that such mismeasurement may have on the business cycle.

The remainder of this paper is organised as follows. Section II addresses the statistical properties of the discrepancy. In Section III we extend on Weale's (1992) methodology. Our findings suggest that private investment has been subject to the most measurement error and consequently if measured correctly private investment is more volatile than existing measures suggest. This coincides with previous results which suggest that actual investment data is not as volatile as theory might suggest [see Guest and McDonald (1995)]. These results are summarised in Section IV. In Section V we demonstrate that if measured correctly, investment may have a significant impact on the nature of the business cycle. Section VI presents our major conclusions.

II  Analysing the Statistical Discrepancy

In recent years there has been a growing concern about the accuracy and reliability of the national accounts [see McDonald (1973,1975), Johnson (1982), Matthews (1984) Lim (1985), Young 1987)]. Claims suggesting the quality of the national accounts have been significantly undermined in recent years are generally supported by the large and volatile statistical discrepancy. Perhaps the volatility of the statistical discrepancy should be a major concern to those who use and interpret the national accounts. It is clear though, that the larger the swings in the statistical discrepancy, the larger are the inconsistencies or quarter-to-quarter growth rates of income and expenditure estimates of GDP over time. The grounds for such concern are primarily two fold; (1) the data source and procedures used to construct missing observations
are unavailable to the statistician from existing data sources [De Leeuw (1990)], and (2) the timing of the recording of transactions [McDonald and Monk (1975)].

II a. Data Sources and Procedures

It is generally costly to collect information frequently, so naturally the ABS makes use of interpolations and extrapolations in order to construct, for example, quarterly observations if it only has annual data. In the case of private consumption the ABS uses household survey data from surveys which it conducts from time to time and constructs estimates of missing observations, both annual and quarterly, by interpolating and/or extrapolating between given sample information. When producing quarterly annual tax estimates of corporate trading enterprise gross operating surplus, the ABS uses its Survey of Operating Surplus developed primarily for this purpose. Although this survey is regarded as a reliable data source, there are instances where other indirect means of obtaining estimates of variables are clearly not so reliable. For example the ABS uses retail turn-over and capital expenditure on dwellings as an activity indicator for the construction of non-farm unincorporated trading enterprise gross operating surplus. In extreme cases where such indirect means are simply not available, the ABS crudely constructs quarterly estimates by interpolating between benchmarks using a linear trend if only annual figures are available.

The use of interpolation techniques are not without their problems. It has been shown [see Milbourne and Bewley (1992)] that a quarterly time series constructed using linear interpolation of annual data will, in a dynamic framework, appear to be Granger caused by the statistical discrepancy. This is because when quarterly data are interpolated from annual data, the quarterly estimates are in fact functions of the annual benchmarks from which they were constructed. Trying to establish causal relationships then is clearly not appropriate.
II b. Timing Issues

The second area of concern relating to the accuracy and reliability of the national accounts are the timing of recording of transactions. Direct observation of the statistical discrepancy series highlights the existence of timing problems. When the recording of transactions are made with some delay, the effects show up in the statistical discrepancy as combinations of volatility and seasonality. In fact, since the mid 1970's the statistical discrepancy has increased significantly in size and volatility and in real terms is highly seasonal. Figure 1 plots the real quarterly statistical discrepancy since September 1959.

This seasonality is consistent with the fact that the measurement of aggregates in the national accounts are subject to significant timing problems, some more than others. Income taxation, for example, is a source of statistical information which is used in the national accounts but is available with a lag. This lag is approximately two years for companies and about one year for individuals, sole traders, partnerships and trusts. Prior to 1978-79 this latter group was also subject to a 2 year lag.

Similar timing inconsistencies are the case for exports and imports. Because the measurement of exports and imports are based on trade data compiled by the Australian Customs Service, the processing of documents does not necessarily coincide with the actual transaction itself. Recently however electronic recording of trade data have significantly reduced timing delays in recording such transactions.

The sensitivity of the timing problem may be illustrated by the following example. Suppose that in June 1999 the recording of a particular transaction ($100m) had been delayed by one quarter. This would affect three quarterly growth rate centred on the quarter at which the delayed transaction is recorded. The measured effects on the growth rates would be -0.09%, 0.19% and -0.1% respectively. The effects on individual components in the national accounts would be much larger [Johnson (1982)].
III Statistical Model

Since by definition, GDP is the market value of goods and services produced in any economy over a period of time, we can define aggregate expenditure as the sum of the smaller expenditure components. Since in the absence of measurement error, GDP\(E) = GDP(I), it must be true that the sum of the smaller income components equal the sum of all the expenditure components. The asterisks defines the true value of the aggregates.

\[
GDP(E)^* = C_p^* + C_g^* + I_p^* + I_g^* + X^* - M^* \\
\equiv W^* + G^{OS^*} + TS^* \\
\equiv GDP(I)^*
\]

where

GDP\(E)^* = the expenditure measure of GDP

GDP\(I)^* = the income measure of GDP

\(C_p^* = Private final consumption expenditure\)

\(C_g^* = Public final consumption expenditure\)

\(I_p^* = Investment (private gross fixed capital expenditure + increases in stocks)\)

\(I_g^* = Public gross fixed capital expenditure\)

\(X^* = Exports of goods and services\)

\(M^* = Imports of goods and services\)

\(W^* = Wages, salaries and supplements\)

\(G^{OS^*} = Gross Operating Surplus\)

\(TS^* = Indirect taxes less subsidies\)

Each component in (1) is subject to measurement error for reasons discussed in Section II. In the national accounts the following expression holds:

\[
GDP(E) = C_p + C_g + I_p + I_g + X - M + SD \\
= W + G^{OS} + TS \\
= GDP(I)
\]

(2)
where SD denotes the statistical discrepancy.

Since the statistical discrepancy is defined as the sum of the measurement error of each component of aggregate demand, we can write (2) as

$$\text{GDP}(E)^* = (C_p + \varepsilon_{cp}) + (C_g + \varepsilon_{cg}) + (I^*_p + \varepsilon_I) + (I^*_g + \varepsilon_K) + (X + \varepsilon_x) - (M + \varepsilon_m)$$

$$= (W + \varepsilon_w) + (G^{OS} + \varepsilon_R) + (TS + \varepsilon_{TS})$$

$$= \text{GDP}(I)^*$$ \hspace{1cm} (3)

where

$$\text{SD} = \varepsilon_{cp} + \varepsilon_{cg} + \varepsilon_{Ip} + \varepsilon_{Ig} + \varepsilon_x - \varepsilon_m + \varepsilon_w + \varepsilon_{G^{OS}} + \varepsilon_{TS}$$

$$C^*_p = C_p + \varepsilon_{cp}; \quad C^*_g = C_g + \varepsilon_{cg}; \quad I^*_p = I_p + \varepsilon_I; \quad I^*_g = I_g + \varepsilon_K; \quad X^* = X + \varepsilon_x;$$

$$M^* = M + \varepsilon_m; \quad W^* = W + \varepsilon_w; \quad G^{OS*} = G^{OS} + \varepsilon_{G^{OS}}; \quad TS^* = TS + \varepsilon_{TS}$$

We apportion SD between $\varepsilon_{cp}, \varepsilon_{cg}, \varepsilon_{Ip}, \varepsilon_{Ig}, \varepsilon_x, \varepsilon_m, \varepsilon_w, \varepsilon_{G^{OS}}$ and $\varepsilon_{TS}$ using a generalisation of Weale (1992)\(^2\). We begin with a (9x1) vector of accounting aggregates $Y$ as measured by the national accounts and another vector $Y^*$, unobservable but true measures of the same aggregates.

$$Y = \begin{bmatrix} C_p \\ C_g \\ I_p \\ I_g \\ X \\ M \\ W \\ G^{OS} \\ TS \end{bmatrix} \quad Y^* = \begin{bmatrix} C^*_p \\ C^*_g \\ I^*_p \\ I^*_g \\ X^* \\ M^* \\ W^* \\ G^{OS*} \\ TS^* \end{bmatrix}$$ \hspace{1cm} (4)

\(^2\) Since the statistical discrepancy is a measure of 'net' error this apportionment is at best an approximation of the truth. It will however give a more accurate relative contribution of each component of aggregate demand on the size of the statistical discrepancy.
Given that the measured values are subject to measurement error, we can write

\[ Y = Y^* + \varepsilon \]  

(5)

We introduce a vector \( K = [K_{11}, K_{12}, K_{13}, K_{14}, K_{15}, K_{16}, K_{17}, K_{18}, K_{19}] \) of accounting constraints such that \( KY^* = 0 \). For our purposes \( K_{11} = K_{12} = K_{13} = K_{14} = K_{15} = 1 \) and \( K_{16} = K_{17} = K_{18} = K_{19} = -1 \) since imports enters positively. This implies, in a sequence of N observations (t=1...N), that the following equality must hold

\[
K Y^* = K_{11} C_p^* + K_{12} C_g^* + K_{13} I_p^* + K_{14} I_g^* + K_{15} X^* + K_{16} M^* + K_{17} W^* + K_{18} G^{05*} + K_{19} TS^* = 0
\]  

(6)

Assuming that [1] \( \varepsilon_t \) and \( Y_t^* \) are uncorrelated, [2] \( \varepsilon_t \) is identically normally distributed with a mean of zero, and [3] \( \varepsilon_t \) are serially independent, we can estimate the true unobserved values of each of the aggregates in (4) by maximising the following log-likelihood function:

\[
L(Y^*, V \mid Y) = -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(V) - \frac{1}{2} (Y - Y^*)^T V^{-1} (Y - Y^*) 
\]  

(8)

subject to the constraint (6). This produces the following result:

\[
\hat{Y}^* = [I - VK^T (KVK^T)^{-1} K]Y 
\]  

(9)

where \( V \) is a \((9\times9)\) unknown variance covariance matrix.

Weale (1992) demonstrates that by probability limits (9) converges to

\[
\hat{Y}^* = [I - D K^T (K^T D K^T)^{-1} K]Y 
\]  

(10)

where \( D \) is a \((9\times9)\) maximum likelihood data covariance matrix and \( I \) is an identity matrix of dimension \((9\times9)^3\).

\[3 \text{ It can also be shown that this result is unaffected by common autocorrelation [see Weale (1992)].}\]
IV Empirical Results

We employ real (at 1989/90 prices) seasonally adjusted data from the domestic production account of the ANA. Our model [eqn 10] is estimated using quarterly national accounts data for the period 1959.3 to 1997.2. A full description of the data is in Appendix A.

In levels the data presented two obstacles. The first obstacle is heteroscedasticity. The presence of heteroscedasticity ensures that the variance of the measurement error will increase as GDP increases over time. A logarithmic transformation of the data has the advantage of reducing the presence of heteroscedasticity by compressing the scale in which the variables are measured. However a using a logarithmic transformation of the data disturbs the accounting constraint given by (6). The second obstacle in non-stationarity. This is generally overcome by first differencing the data.

To ensure that our model produces consistent results with Weale's simplified model, we require an alternative data transformation. For the smaller expenditure aggregates, namely $C_p, C_g, I_p, I_g, X$ and $M$, we take first differences as a proportion of GDP(E). For the smaller income aggregates, namely $W, G^o$ and TS, we take first differences as a proportion of GDP(I). These transformations ensure that [1] the accounting constraint (6) is not disturbed, [2] the data is stationary. Testing for the presence non-stationarity we find that in levels all the variables exhibit evidence of a unit root but are stationary under the new transformation. These results are reported in Table 1: and [3] the sum of the proportion, $\delta$, of the statistical discrepancy contributed by each of the smaller national accounting aggregates sum to the proportions contributed by the larger aggregates, namely GDP(I) and GDP(E), that is,

---

4 Weale (1992) used logarithmic first differenced data since the methodology is unaffected using this particular data transformation in a 2 variable case.
\[ \delta_C \left( \frac{\Delta C}{Y(E)_{t-1}} \right) + \delta_I \left( \frac{\Delta I}{Y(E)_{t-1}} \right) + \delta_G \left( \frac{\Delta G}{Y(E)_{t-1}} \right) + \delta_{NX} \left( \frac{\Delta NX}{Y(E)_{t-1}} \right) = \delta_{Y(E)} \left( \frac{\Delta GDP(E)_t}{Y(E)_{t-1}} \right) \]

and

\[ \delta_w \left( \frac{\Delta W_t}{GDP(I)_{t-1}} \right) + \delta_{G^{os}} \left( \frac{\Delta G^{os}}{GDP(I)_{t-1}} \right) + \delta_{TS} \left( \frac{\Delta TS_t}{GDP(I)_{t-1}} \right) = \delta_{Y(I)} \left( \frac{\Delta GDP(I)_t}{GDP(I)_{t-1}} \right) \]

such that \( \delta_C + \delta_I + \delta_G + \delta_{NX} = \delta_{Y(E)} \); \( \delta_w + \delta_{G^{os}} + \delta_{TS} = \delta_{Y(I)} \) and

where \( C = C_p \); \( G = C_g + I_i \) and \( NX = X - M \).

We are now in a position to construct maximum likelihood estimates for the true but unobservable values of national output. We repeat Weale's approach and apportion the statistical discrepancy between each of the measured values of GDP(E) and GDP(I)\(^5\). The proportion of the statistical discrepancy attributed to each of these two inconsistent measures of national output are given in Table 2. Also in Table 2 is a measure of how sensitive changes in sample size are on the measurement of these proportions.

Table 2 suggests that the statistical discrepancy is predominantly unmeasured aggregate expenditure. Approximately 90\% of the statistical discrepancy is the result of measurement error in GDP(E) with 10\% accounted for by mis-measurement in aggregate income\(^6\). These results appear relatively robust to changes in sample size. Varying the sample size has no significant effect on these results.

In Table 3 we present results for the generalised model. The proportion of the statistical discrepancy contributed by each of the smaller national accounting aggregates are presented in column 1 of Table 3. As expected the sum of the contributions from \( C_p, C_g, I_p, I_i, X \) and \( M \) sum in absolute value to the contributions for GDP(E). Similarly the contributions from \( W, G^{os} \) and IT sum in absolute value to

---

\(^5\) To ensure that both our model and Weale's model produce consistent results, GDP(I) and GDP(E) are measured in quarterly growth rates. This is consistent with Weale's logarithmic first differencing of the data.

\(^6\) Weale (1992) also finds the expenditure measure contributes most to the size of the statistical discrepancy in the United States.
the contribution for GDP(I). The proportion of the statistical discrepancy contributed by GDP(E) and GDP(I) are those reported in Table 2.

These results demonstrate, as might be expected, that government data is less likely to be influenced by measurement error than is non-government data. Public consumption expenditure was found to contribute approximately 4% of total measurement error in the national accounts. Trade data appears to do relatively well with imports contributing the least to the measurement error (2%) and exports contributing 4% of total measurement error.

Private final consumption and private investment expenditure do not perform so well. Private consumption expenditure contributes 6% to the total measurement error while private investment expenditure contributes a staggering 69% of total measurement error. The extent of the measurement error in investment\(^7\) (74%) may have significant impact of the volatility of the business cycle (to be discussed below) particularly if the volatility of correctly measured private investment is greater than the volatility of existing measures.

In columns 2 and 3 of Table 3 an exercise is undertaken to reinforce the robustness of these results. In column 2 the sum of \(C_p, I_p, G = C_e + I_e\), \(X\) and \(M\) in absolute value sum to GDP(E) as does the sum of \(C_p, I, G\) and \(NX = X - M\) [column 3]. From either disaggregation (columns 1, 2 or 3) private investment expenditure contributes to approximately 74% of the size and variation of the statistical discrepancy. It is to the implications of such mis-measurement, particularly private investment expenditure, that we now turn.

V Mismeasurement of Investment

There have been a number of attempts [see Milbourne and Bewley (1992), McKibbin and Morling (1989) and Gregory (1989)] to determine which aggregates in the national accounts have contributed most to the size and variability of the statistical discrepancy.

---

\(^7\) The sum of private and public gross fixed capital expenditure and increases in stocks.
discrepancy. Gregory (1989) takes the view that private sector saving-investment imbalances may explain most of the measurement error in the national accounts.

McKibbin and Morling (1989) argue that the statistical discrepancy is unmeasured consumption expenditure. This argument is mistakenly premised on a negative correlation found to exist between the statistical discrepancy and consumption. It is possible to demonstrate that such a correlation naturally exists\(^8\). Assuming the measurement error in consumption, \(e_c (= e_p + e_g)\), is uncorrelated with the statistical discrepancy \((e_{sd})\), the covariance between the two may be written as follows:

\[
\sigma_{sd,c} = E[SD - E[SD]] \cdot [C - E(C)] \\
= E[SD \times (C^* + e_c - E[C])] \\
= -\sigma^2_e 
\]

since \(E[SD] = 0\), \(e_c\) is white noise and \(SD = -e_c - e_i - e_g - e_x + e_m\).

Similarly, it is possible to show that the covariance between the remaining expenditure components and the statistical discrepancy are:

\[
\sigma_{sd,I} = -\sigma^2_{e_i}; \quad \sigma_{sd,G} = -\sigma^2_{e_g}; \quad \sigma_{sd,X} = -\sigma^2_{e_x}; \quad \sigma_{sd,M} = \sigma^2_{e_m}
\]

Therefore, as long as \(e\) is white noise the covariance of the statistical discrepancy with private and public consumption and investment expenditure and exports should be negative. The covariance of the statistical discrepancy with imports should be positive. Table 4 presents the covariances between each of the aggregates in the ANA and the statistical discrepancy. As expected the covariances have the right sign.

Milbourne and Bewley (1992) more specifically argue that a significant proportion of measurement error in the national accounts arises from private sector investment expenditure. This appears to conform with our results that private investment expenditure is significantly mis-measured. There are several implications for this. First, changes in investment usually conveys valuable information about the future movements in the economy and measurement error is only likely to bias such

\(^8\) See Milbourne and Bewley (1992)
information. Second, changes in investment have a significant impact on the movements in national output and hence may suggest that the business cycle is more volatile than is actually reported.

With over 74% of the statistical discrepancy attributed to the mis-measurement of private and public investment\(^9\), we construct a new measure of investment, \(I^*\) with equation (9). In Figure 2 we plot this new measure of investment against available estimates from 1988:4 to 1997:4. There are clearly three distinct phases in Figure 2. The first phase, (1988:4 to 1991:1), suggests that investment has been consistently under-reported. The second phase, (1991:1 to 1994:3), investment was over-reported, with the exception of June and September of 1993. The final phase, (1994:3 to 1997:2), investment was again consistently under-reported. This suggest that investment is under-reported in boom periods and over-reported in periods of slow economic growth or during recessions.

In Figure 3 we plot the growth rate of \(I^*\) and investment data as reported in the national accounts. We find that \(I^*\) has a greater variance in its growth rate \((\sigma^2 = 39)\) than does the existing measure of investment \((\sigma^2 = 30)\). This has a direct policy implication. Since the presence of non-stationarity in investment means that theories of investment are tested in first difference, the greater the volatility in \(I^*\) may substantially change existing policy implications based on incorrectly measured investment data.

Blinder (1981) and Blinder and Maccini (1991) have suggested that in periods of recession, falls in investment account for the bulk of the decline in GDP. In Table 5 we date the growth cycle using the Bry-Boschan (1971) business cycle dating procedure\(^10\) for the new measure of investment, \(I^*\), and GDP(I). Both \(I^*\) and GDP(I) share similar turning points in the growth cycle which suggests that investment may have a significant impact on the short run growth of GDP. Investment, \(I^*\), is a noisy time series and consequently the average peak-to-peak (37 months) and the average trough-to-trough (36 months) durations are shorter than those for GDP, which are 52

---

\(^9\) Although only 5% of the total measurement error is accounted for by public investment expenditure.

\(^10\) This procedure is based on the well known NBER business cycle dating methodology.
and 50 months respectively. In Figure 4 we plot the business cycle components for these two data series. It appears that for Australia the cyclical movements in investment, \( I' \), share a similar cyclical patterns present in output. This implies that any improvement in the measurement of investment which affects its variance may significantly impact on the movement of measured national output and hence the business cycle.

Table 6 shows the contributions of investment, including \( I' \), to GDP growth for the Australian business cycle since 1960. The first two columns of Table 6 identify the peaks and troughs of the business cycle. The third column identifies the peak-to-trough change in GDP and the fourth and fifth columns, the contributions to GDP growth from present measures of investment and the new measure of investment, \( I' \). The sixth, seventh and eight columns show similar figures for the first year of recovery.

The first notable observation from Table 6 is that the new measure of investment, \( I' \), contributes much more to GDP growth than implied by existing measures of investment. For example, in 1960/61 the contributions to GDP from existing measures of investment was three times the fall in GDP while the contributions from \( I' \) was nearly four times the fall in GDP. During the first year of recovery the differences between the two measures of investment were not so great. \( I' \) contributed 0.14% more to GDP growth than the existing measures of investment. This suggests that investment has a greater effect during the downswing of the business cycle than it does during a recovery. Throughout 1990/91, \( I' \) contributed 1.67% more to GDP growth than implied from existing official statistics on investment. During the first year of recovery however, the contributions to GDP growth from the existing measures of investment was only one-quarter of the growth in GDP, while the contribution from \( I' \) was only 7% of the growth of GDP. Whichever point in the business cycle one examines, the new measure of investment, appears to contribute significantly to changes in output, hence the business cycle, particularly so during a downswing.
VI Conclusion

The growth in the statistical discrepancy particularly since the mid 1980’s has prompted a number of researchers to investigate the components of aggregate demand likely to have contributed most to its size and variance. Overwhelming evidence seems to suggest that private investment expenditure contributes significantly to measurement errors in the national accounts. Consistent with expectations, it was found that public data appears to be measured more accurately simply because it is determined directly by government and good records of this data exist.

Having estimated that private investment expenditure contributes three-quarters of the total measurement error in the national accounts, an interesting result came to light. First, the volatility of an error corrected investment series is much larger than the variance of the existing measure of investment. Ultimately this may have a significant implication for testing existing theories of investment. Second, and equally interesting, the new measure of investment has a significant impact on the nature of the business cycle in Australia, namely that it increases business cycle volatility.

It is imperative therefore that policy makers take seriously the implications of measurement error in the national accounts while investigating avenues to improve the quality of variables measured.
Appendix A: Data Source and Description

We employ the following real (at 1989/90 prices) seasonally adjusted data - (Source: ABS Time Series (TS): Table 5206-22: Domestic Production Account (DPA)-Seasonally Adjusted) using the implicit GDP(E) deflator. It is constructed as the ratio of GDP (exp based, current prices; Source: ABS TS: DPA - Table 5206.23) and GDP (exp based, 1989/90 prices; Source: ABS TS: Measures of GDP - Table 5206-1). DX Database identifiers in brackets.

\[ \text{C}_p = \text{Private Final Consumption Expenditure} \quad \text{[NADQ.AC#PH#99FCE]} \]
\[ \text{C}_g = \text{Government Final Consumption} \quad \text{[NADQ.AC#GG#99FCE]} \]
\[ \text{I}_p = \text{Investment. This is made up of two categories:} \]

1. \textit{Private Gross Fixed Capital Expenditure} - constructed as the sum of the following four categories:
   a. Dwellings \text{[NADQ.AC#P##99GFC_DWL]}
   b. Non-dwelling Construction \text{[NADQ.AC#P##99GFC_NDC]}
   c. Equipment \text{[NADQ.AC#P##99GFC_EQP]}
   d. Real Estate Transfer Expenses \text{[NADQ.AC#P##99GFC_RET]}

2. \textit{Increases in Stocks} - the sum of the following four categories:
   a. Private Non-farm Stocks \text{[NADQ.AC#P##98IST]}
   b. Farm Stocks \text{[NADQ.AC#IS_FAR#]}
   c. Public Marketing Authority Stocks \text{[NADQ.AC#IS_PMA#]}
   d. Other Public Authority Stocks \text{[NADQ.AC#IS_OPA#]}

\[ \text{I}_g = \text{Public Gross Fixed Capital Expenditure. This is made up of two categories:} \]

1. Public Enterprise Gross Fixed Capital Expenditure \text{[NADQ.AC#GE#99GFC]}
2. General Government Gross Fixed Capital Exp \text{[NADQ.AC#GG#99GFC]}

X = Exports of Goods and Services \text{[NBDQ.AC_XGS#]}
M = Imports of Goods and Services \text{[NBDQ.AC_MGS#]}
GDP(E) = Gross Domestic Product: Expenditure Measure
SD = Statistical Discrepancy: Difference between the real seasonally adjusted income measure and the exp measure of GDP \text{[NODQ.AL_STAT_DIS]}
GDP(I) = Gross Domestic Product: Income Measure \text{[NODQ.AC_GDP]}
W = Wages, Salaries and Supplements \text{[NWDQ.ACW_#T_99WS]}
G^{05} = Gross Operating Surplus: The sum of the following six categories:
   a. Gross Operating Surplus: Private Trading Corporate Enterprises \text{[NIDO.AC_GOS_TEAA]}
   b. Gross Operating Surplus: Private Trading Unincorporated Enterprises \text{[NIDO.AC_GOS_UNIC]}
   c. Gross Operating Surplus: Private Trading Enterprises: Dwelling owned by persons \text{[NIDO.AC_GOS_DWEL]}
   d. Gross Operating Surplus: Public Trading Enterprises \text{[NIDO.AC_GOS_PUTE]}
   e. Gross Operating Surplus: Financial Enterprises (less imputed bank service changes) \text{[NAD.AC_GOS_FELC]}
   f. Gross Operating Surplus: General Government \text{[NADQ.UC#GG#99CFC]}

TS = Indirect Taxes less Subsidies \text{[NIDQ.AC_ITX_LSUB]}

16
## Tables

### Table 1 - Dickey-Fuller Unit Root Tests

<table>
<thead>
<tr>
<th>Variable</th>
<th>Levels</th>
<th>First Difference</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Constants</td>
<td>No Trend</td>
</tr>
<tr>
<td></td>
<td>$\alpha(1)=0$</td>
<td>$\alpha(1)=\alpha(2)=0$</td>
</tr>
<tr>
<td>$Y_t$</td>
<td>$q$</td>
<td>$\nu$</td>
</tr>
<tr>
<td>$C_p$</td>
<td>6</td>
<td>3.26</td>
</tr>
<tr>
<td>$C_T$</td>
<td>3</td>
<td>-0.05</td>
</tr>
<tr>
<td>$I_p$</td>
<td>10</td>
<td>-0.80</td>
</tr>
<tr>
<td>$I_g$</td>
<td>1</td>
<td>-2.22</td>
</tr>
<tr>
<td>$X$</td>
<td>9</td>
<td>2.99</td>
</tr>
<tr>
<td>$M$</td>
<td>12</td>
<td>2.15</td>
</tr>
<tr>
<td>$W$</td>
<td>0</td>
<td>0.44</td>
</tr>
<tr>
<td>$G^{\text{GDP}}$</td>
<td>0</td>
<td>0.42</td>
</tr>
<tr>
<td>$TS$</td>
<td>11</td>
<td>1.52</td>
</tr>
<tr>
<td>GDP(I)</td>
<td>4</td>
<td>1.39</td>
</tr>
<tr>
<td>GDP(E)</td>
<td>2</td>
<td>1.48</td>
</tr>
</tbody>
</table>

Notes:

a) Columns 3 and 6 have the following Dickey-Fuller specifications:

$$\Delta Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \sum_{i=1}^{q} \beta_i \Delta Y_{t-i} + \epsilon_t$$

and Columns 4 and 7 have the following specifications (where $t$ = time trend)

$$\Delta Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 t + \sum_{i=1}^{q} \beta_i \Delta Y_{t-i} + \epsilon_t$$

b) Null hypotheses are found at the head of each column. $\alpha(1)=0$ in columns 3 and 6 are $t$-tests and in columns 4 and 7, $\alpha(1)=\alpha(2)=0$ are unit root tests with non-zero drift (F-test $\Phi_3$). The critical $t$-statistic for columns 3 and 6 is -2.57, and the critical F-test $\Phi_3$ for columns 4 and 7 is 5.34.

c) $d$ and $q$ were chosen as the highest lag from the autocorrelation function of the first differenced series at the 95% confidence interval.

### Table 2 - Proportion of the Statistical Discrepancy Contributed by GDP(I) and GDP(E)

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\delta$</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
<th>30</th>
<th>35</th>
</tr>
</thead>
<tbody>
<tr>
<td>GDP(E)</td>
<td>0.90</td>
<td>0.91</td>
<td>0.91</td>
<td>0.90</td>
<td>0.91</td>
<td>0.92</td>
<td>0.92</td>
<td>0.93</td>
</tr>
<tr>
<td>GDP(I)</td>
<td>0.10</td>
<td>0.09</td>
<td>0.09</td>
<td>0.10</td>
<td>0.09</td>
<td>0.08</td>
<td>0.08</td>
<td>0.07</td>
</tr>
</tbody>
</table>

Notes: $\delta$ denotes the proportion of the total measurement error attributed to either GDP(E) or GDP(I). Columns 3 to 9 represent the same proportion except the number of end-point observations dropped from the data set is denoted at the head of each column.
### Table 3 – Proportion of the Statistical Discrepancy Contributed by the Components of Aggregate Demand.

<table>
<thead>
<tr>
<th>Variable (1)</th>
<th>δ (2)</th>
<th>Variable (3)</th>
<th>δ (4)</th>
<th>Variable (5)</th>
<th>δ (6)</th>
<th>Variable (7)</th>
<th>δ (8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$C_p$</td>
<td>0.06</td>
<td>$C_p$</td>
<td>0.06</td>
<td>$C_p$</td>
<td>0.06</td>
<td>$C_p$</td>
<td>0.06</td>
</tr>
<tr>
<td>$C_g$</td>
<td>0.04</td>
<td>$I_p$</td>
<td>0.69</td>
<td>$I_p$</td>
<td>0.69</td>
<td>$I_p$</td>
<td>0.69</td>
</tr>
<tr>
<td>$I_g$</td>
<td>0.69</td>
<td>$C_g + I_g$</td>
<td>0.09</td>
<td>$C_g + I_g$</td>
<td>0.09</td>
<td>$C_g + I_g$</td>
<td>0.09</td>
</tr>
<tr>
<td>$I_s$</td>
<td>0.05</td>
<td>$X$</td>
<td>0.04</td>
<td>$X - M$</td>
<td>0.06</td>
<td>$X - M$</td>
<td>0.06</td>
</tr>
<tr>
<td>$X$</td>
<td>0.04</td>
<td>$M$</td>
<td>0.02</td>
<td>$W$</td>
<td>0.03</td>
<td>$G^{me}$</td>
<td>0.06</td>
</tr>
<tr>
<td>$M$</td>
<td>0.02</td>
<td>$W$</td>
<td>0.03</td>
<td>$G^{me}$</td>
<td>0.06</td>
<td>$W$</td>
<td>0.01</td>
</tr>
<tr>
<td>$G^{me}$</td>
<td>0.06</td>
<td>$W$</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TS</td>
<td>0.01</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 4 – Covariance of the Statistical Discrepancy with the Expenditure Components of the National Accounts.

<table>
<thead>
<tr>
<th></th>
<th>Covariance</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$C_p$</td>
<td>$C_g$</td>
<td>$I_p$</td>
<td>$I_s$</td>
<td>$X$</td>
<td>$M$</td>
<td></td>
</tr>
<tr>
<td>SD</td>
<td>-0.47E-04</td>
<td>-0.35E-04</td>
<td>-0.38E-04</td>
<td>-0.20E-04</td>
<td>-0.65E-05</td>
<td>0.17E-05</td>
<td></td>
</tr>
</tbody>
</table>

**Notes:** The covariances are measured between each of the aggregate demand components as a proportion of GDP and the statistical discrepancy as a proportion of GDP.
Table 5 - Growth Cycle Turning Points

<table>
<thead>
<tr>
<th>Peak</th>
<th>Trough</th>
<th>Peak</th>
<th>Trough</th>
</tr>
</thead>
<tbody>
<tr>
<td>1963:06</td>
<td>1963:12</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>1967:06</td>
<td>1968:06</td>
</tr>
<tr>
<td>1971:06</td>
<td>1972:12</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1977:09</td>
<td>1978:03</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>1982:03</td>
<td>1983:09</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Average P-P: 37 months
Average T-T: 36 months
Average P-T: 15 months
Average T-P: 22 months

Notes: Linear interpolation of the data was required to generate monthly observations of the quarterly data in order to implement the set of Bry-Boschahn business cycle dating procedures.

Table 6 - Contributions of Investment to GDP Growth

<table>
<thead>
<tr>
<th>Peak</th>
<th>Trough</th>
<th>%Δ GDP</th>
<th>I</th>
<th>I'</th>
<th>%Δ GDP</th>
<th>I</th>
<th>I'</th>
</tr>
</thead>
<tbody>
<tr>
<td>1974:03</td>
<td>1977:12</td>
<td>-0.50</td>
<td>-0.46</td>
<td>0.10</td>
<td>2.99</td>
<td>-4.66</td>
<td>-6.55</td>
</tr>
<tr>
<td>1981:06</td>
<td>1986:03</td>
<td>-2.34</td>
<td>-5.53</td>
<td>-6.42</td>
<td>8.41</td>
<td>4.03</td>
<td>6.03</td>
</tr>
<tr>
<td>1990:03</td>
<td>1991:06</td>
<td>-2.94</td>
<td>-4.71</td>
<td>-6.37</td>
<td>2.53</td>
<td>0.63</td>
<td>0.17</td>
</tr>
</tbody>
</table>
Figures

Figure 1 – *Real (at 1989/90 prices) Statistical Discrepancy*

Figure 2 – *Measures of Investment: New (I*) and Existing*
Figure 3 – Measured Growth Rates of Investment: New ($I^*$) and Existing

Figure 4 – Business Cycle Components of GDP and the New Measure of Investment ($I^*$)
References


